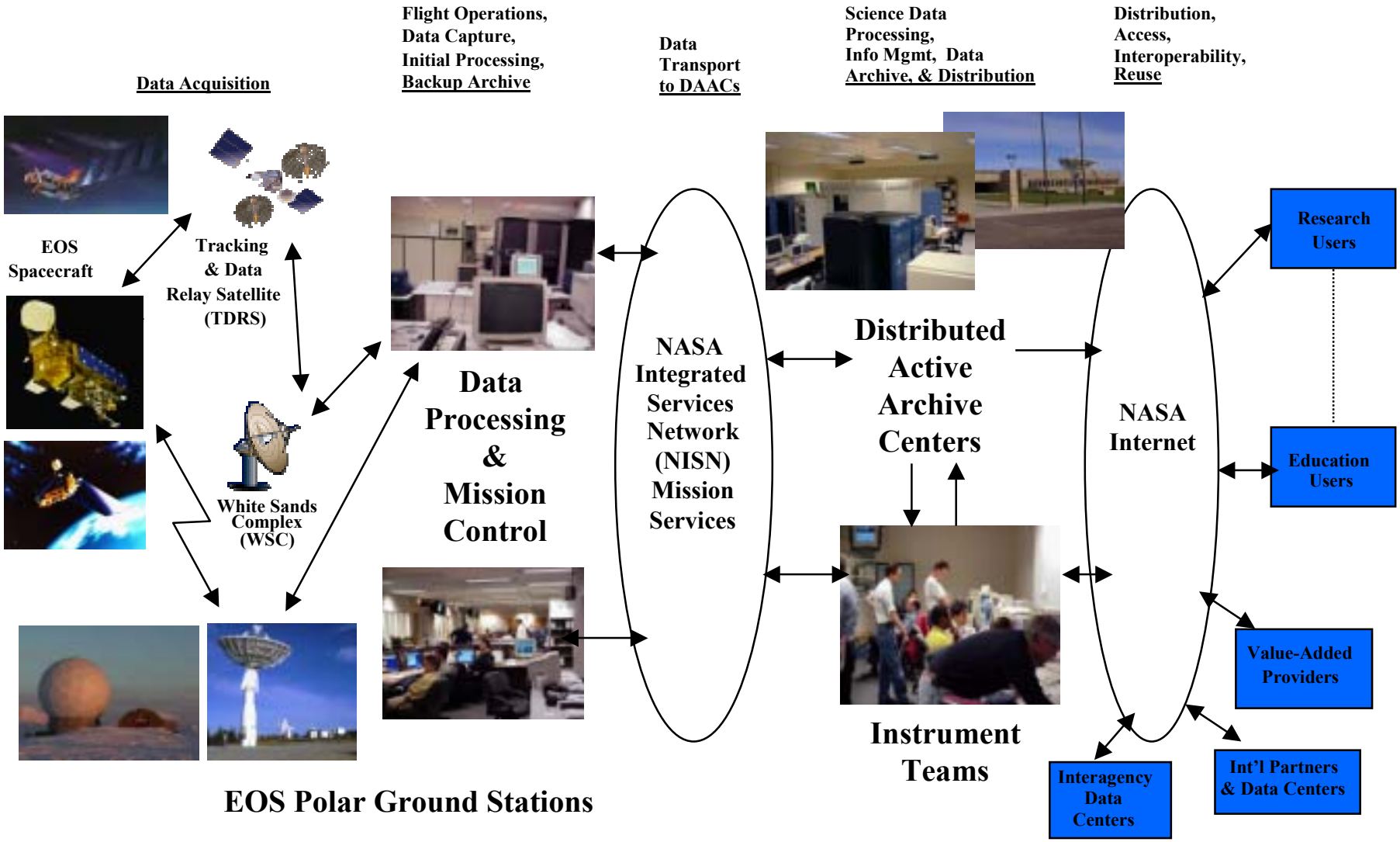
ESE Data Systems and Services for the Next Decade: Strategic Evolution for Effective Information Delivery and Utilization

March 18, 2003

Martha Maiden
Program Executive, Data Systems
Office of Earth Science

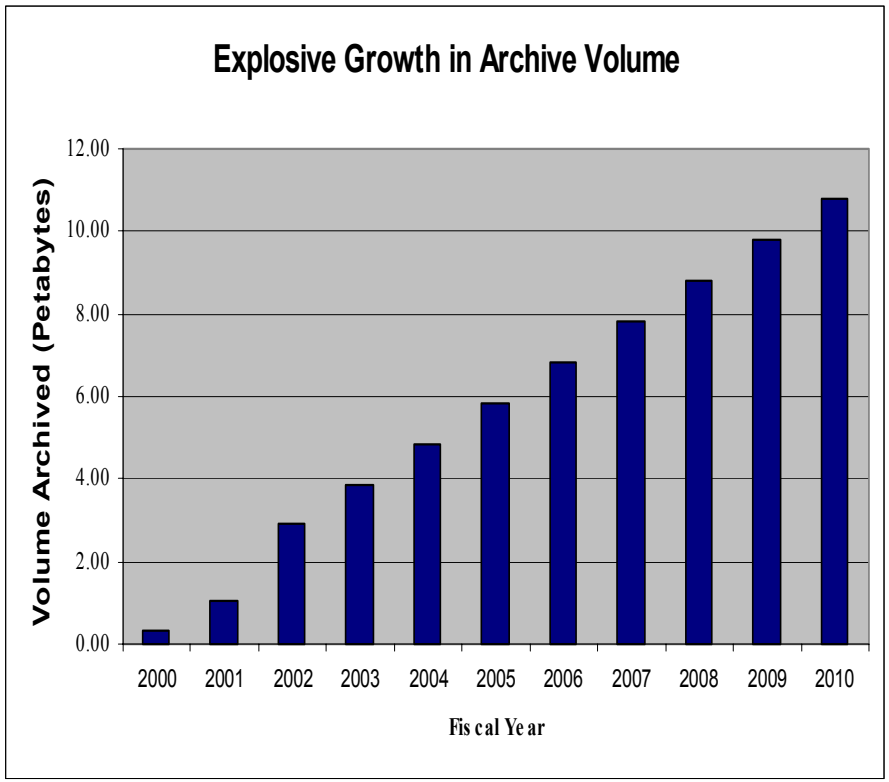
EOSDIS Context Diagram

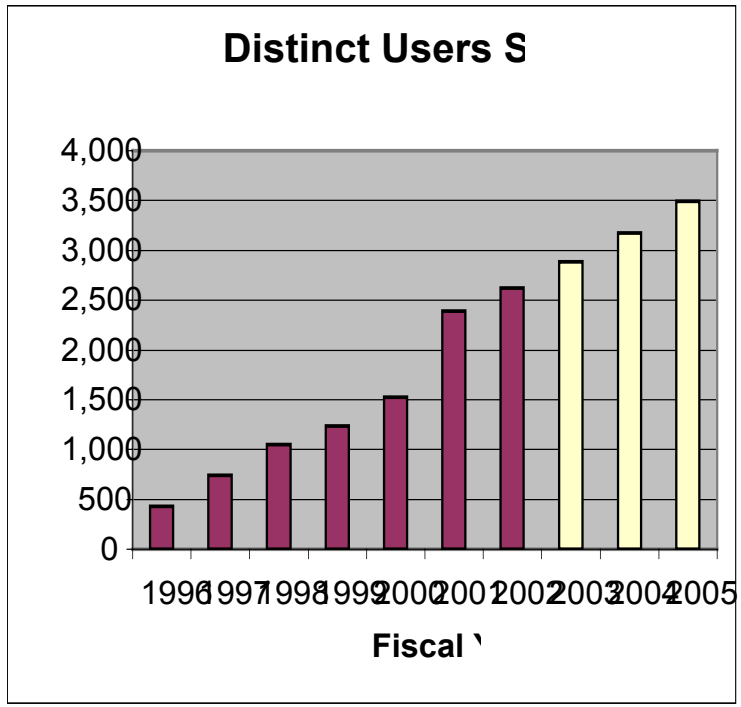


NASA is Meeting a Growing Demand for Earth Science Data and Information

Ingesting, processing, and archiving an unprecedented volume of climate and Earth science data.

NASA is benchmarking capabilities and processes for handling the capacities for future operational needs (e.g., NPOESS).



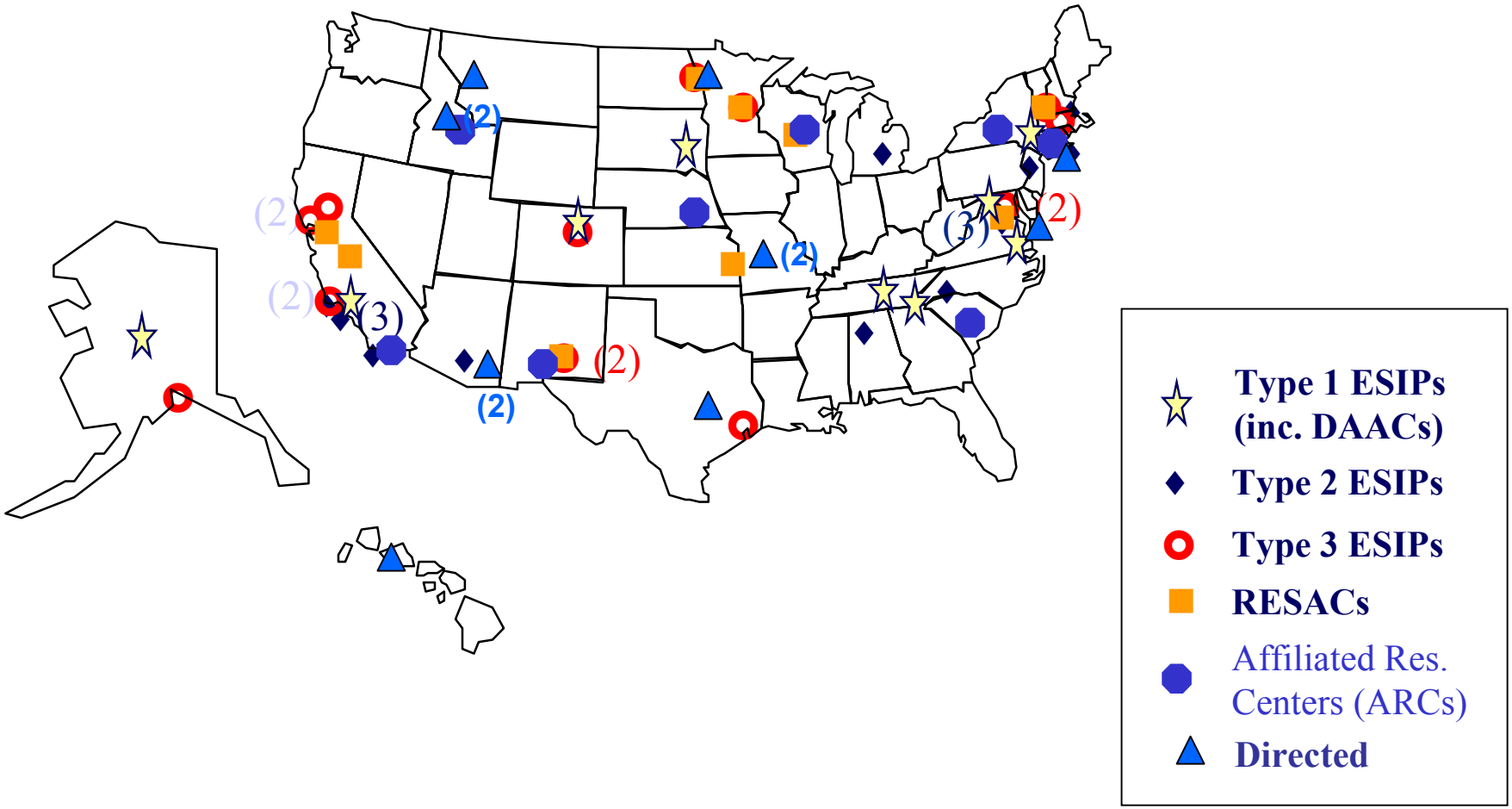


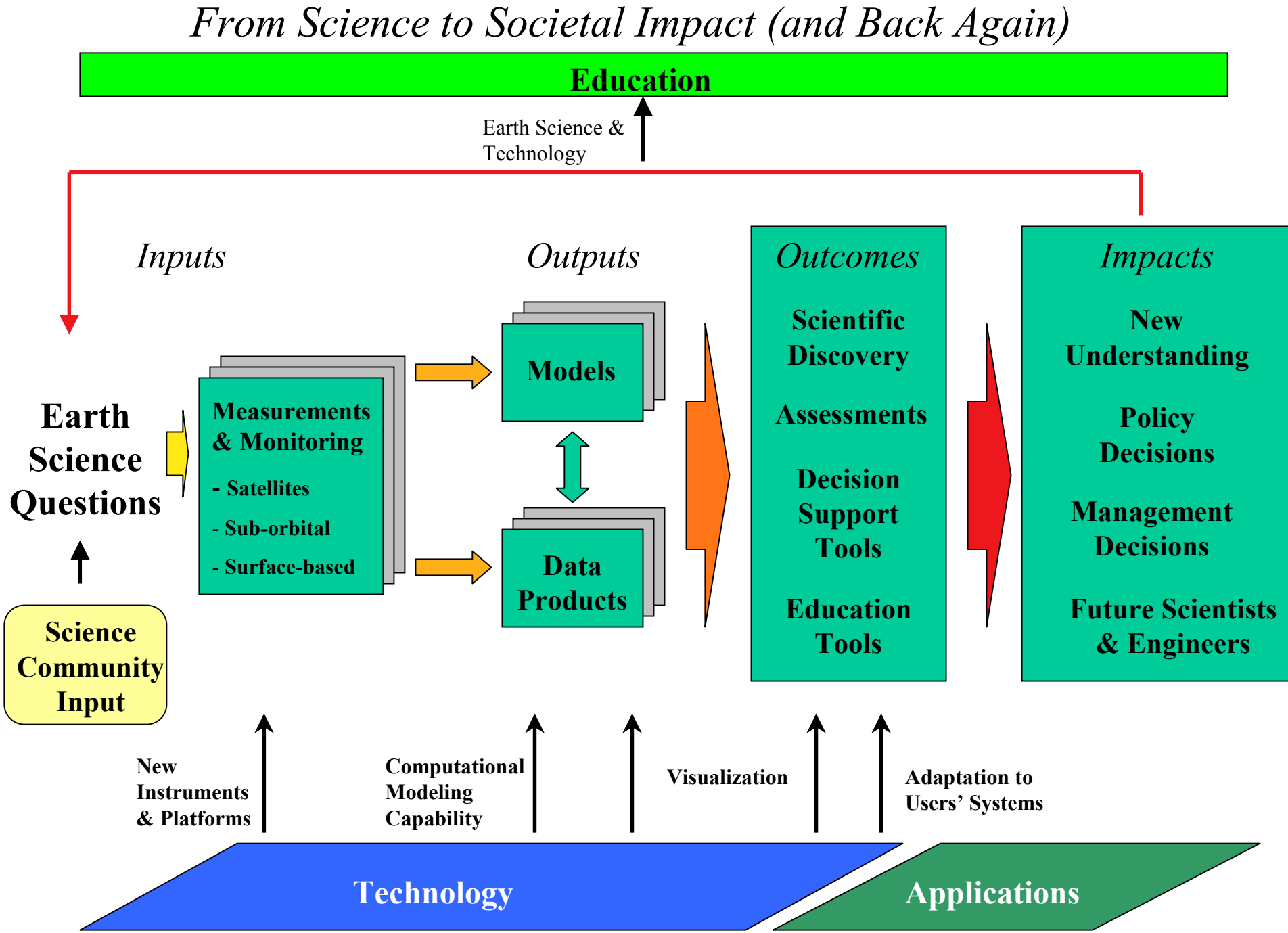
NASA provides access to Earth system science data, information, and services to millions of unique users.

Over the next decade, NASA will ensure the timely delivery of Earth Science information at an affordable cost by evolving to a more open, distributed set of data systems and service providers.

ESE Data Center Locations

ESE supports 68 data centers (some of which at the same location), widely distributed geographically. Additional data centers, including NOAA's NCDC and Unidata, are networked through Membership in the ESIP Federation.




From Science to Societal Impact (and Back Again)
Education
Earth Science & Technology
Inputs
Outputs
Outcomes
Impacts
Earth Science Questions
Measurements & Monitoring
- Satellites
- Sub-orbital
- Surface-based
Models
Data Products
Scientific Discovery
Assessments
Decision Support Tools
Education Tools
New Understanding
Policy Decisions
Management Decisions
Future Scientists & Engineers
Science Community Input
New Instruments & Platforms
Computational Modeling Capability
Visualization
Adaptation to Users' Systems
Technology
Applications

#

How effectively can evolved data and information systems and services support the community to answer Earth science thematic questions?

Goal: Support ESE Goals with a common framework for Distributed Data Information Systems and Services

ESE AA Approval
NRC Review

Early DRAFT

NASA
Joint
Unfunded

Data and Service Accessibility and Utility

Interoperability achieved to enable Climate Change Research Initiative (CCRI)

Agreed standards and protocols, and distributed services provide"interuse" of multi-source data and information

Interface standards automate data flow to models

*Infusion of High-End Computing and Modeling Technologies

Research Groups to Enable Systematic Measurements (CDRs) in place

Institutional expertise are "smart buyers", buttressed by competition for science needs through AOs, NRAs, REASoN CANs

Integrated Precipitation Processing System

SEEDS Prototype Precipitation Processing System supports concepts of reuse, evolution, and thematic architecture

Earth Science Enterprise Data and Information Systems and Services Evolve

REASoN Cooperative Agreements provide science, applications, and education solutions for science themes, competed on 2 1/2 year timescale for 5 year projects.

Strategic Evolution of ESE Data Systems (SEEDS) Formulation

Strategic Framework and Coordinating Activity for Evolution of ESE network of data systems and service providers over the next decade

Strategically evolve ESE data and information systems and services to support science thematic questions.

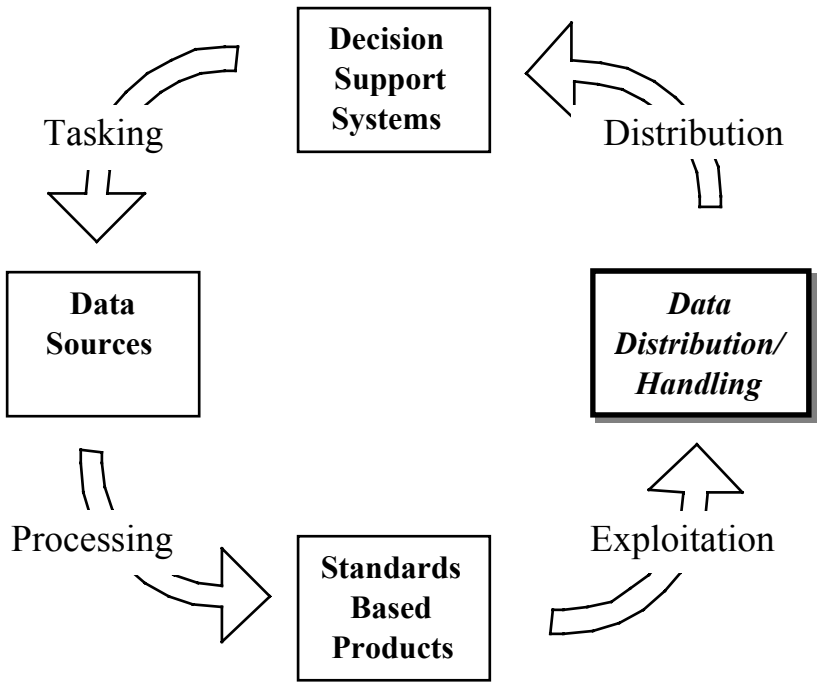
2002: EOSDIS fully developed and supporting two petabytes of data; EOSDIS Maintenance and Development solicited and planning on-going for next decade of systems

Provides Level 0-3 standard product processing, archive, and distribution for EOS Missions: Landsat, Terra, Aqua, Aura, Icesat, SORCE

Systematic Measurements and Fused/"bundled" Data Sets for Interdisciplinary or Process Studies and Models

2002 2004 2006 2008 2010 2012 2014 2015

The ESE Information Cycle



ESE Information Cycle

Typical development cycle for information from user requirement (at the top of the figure) through decision support system. The Data Systems provide Processing (SIPS), Standard Based Products housed in the DAACs and their utilization. REASoNs is a means to complete the cycle by benchmarking solutions network assimilation into Decision Support tools.

REASoN CAN Objective

- The CAN solicits proposals that will afford ***solutions*** for utilization of NASA assets and capabilities by:
  - providing data products and/or information systems and services capabilities in support of the goals and objectives of the research, applications, and education strategies of NASA's Earth Science Enterprise (ESE);
  - developing, where necessary, advanced data systems technologies integrated into a project (solution) that addresses the above objectives.
  - applying principles from the Strategic Evolution of ESE Data Systems (SEEDS) regarding community involvement, product life cycle planning, and standards and interfaces for interoperability and exchange of data and information;
  - supporting ongoing SEEDS efforts through Working Groups for Standards and Interfaces, Technology Infusion, Architecture and Reuse, and Metrics Planning and Reporting;
  - contributing to benchmarking solutions that serve society through integration of Earth science measurements, models and decision support systems.

CAN Objective (cont.)

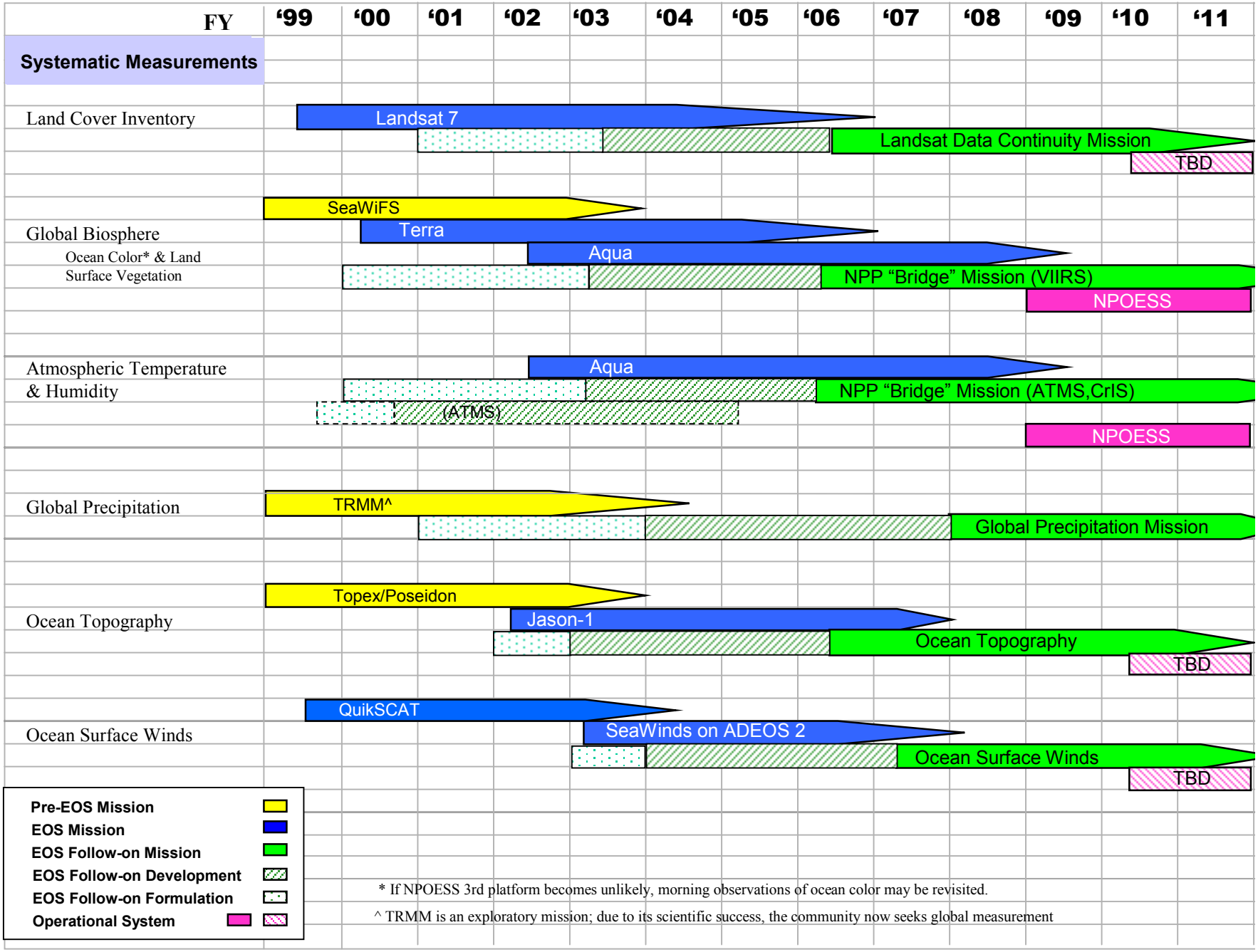
Projects supported by this CAN will provide data and data products and/or information systems and services capabilities to:

(*Research*) Improve accessibility by the NASA science community to, and accuracy of: a) data and data products, including selected geophysical parameters of Earth observations constructed from multiple sources; and, b) efforts that more effectively integrate and fuse sources for geophysical parameters that may not be directly observed;

(*Applications*) Provide data products and tools for resource management and policy decision support in applications of national importance, and provide decision makers with interactive access to dynamically updated knowledge of the Earth system; and,

(*Education*) Address needs of the educational community particularly with respect to timely and ready access to Earth and environmental data to promote math, science and geography in K-12 education, and earth system science in graduate and post graduate education.

ESE Systematic Measurements & Mission Summary

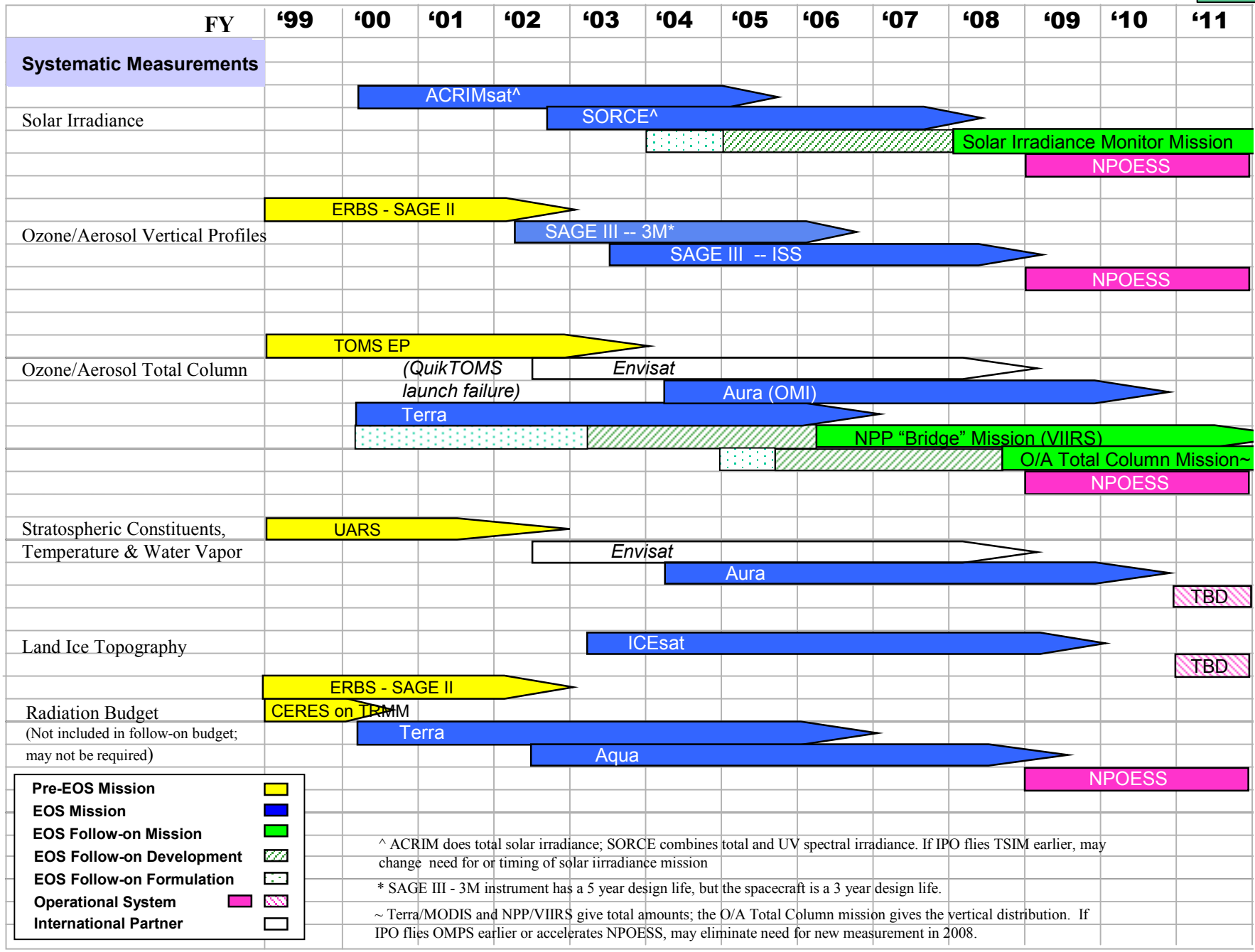


* If NPOESS 3rd platform becomes unlikely, morning observations of ocean color may be revisited.

^ TRMM is an exploratory mission; due to its scientific success, the community now seeks global measurement

ESE Systematic Measurements & Mission Summary

Roadmap



^ ACRIM does total solar irradiance; SORCE combines total and UV spectral irradiance. If IPO flies TSIM earlier, may change need for or timing of solar iirradiance mission

* SAGE III - 3M instrument has a 5 year design life, but the spacecraft is a 3 year design life.

~ Terra/MODIS and NPP/VIIRS give total amounts; the O/A Total Column mission gives the vertical distribution. If IPO flies OMPS earlier or accelerates NPOESS, may eliminate need for new measurement in 2008.

Precipitation Processing System (SEEDS Prototype)

Data Production System:

Closely working federation of national and international partners

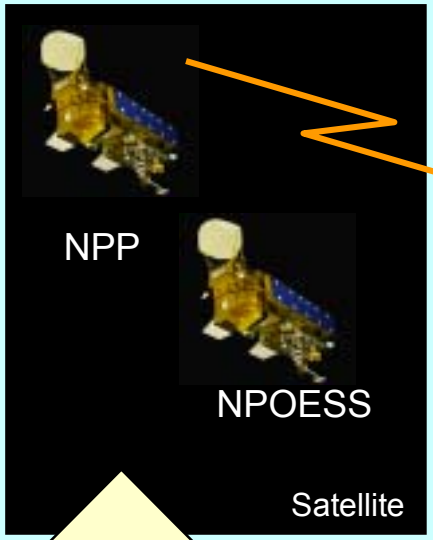
Global Precipitation Measurement Mission Architectural Guidelines

- Don’t view as a single-point mission
- Conceived as a rolling-wave of capabilities
  - Satellites/instruments added, deleted, replaced as required
  - Partners added, deleted, replaced as required
  - Data streams added, deleted, replaced as required
- Data system scalable to handle rolling-wave mission concept
  - Focused on precipitation question rather than a specific mission
  - Minimum and definable costs for increase, replacement, deletion
  - No built-in software and architecture limitations on scalability
  - Ability of other partners to tie into data and services

Precipitation Processing System Approach

- Architecture based (Stocker – Chief Architect)
- Large Scale Reuse from TRMM
  - Entire Architecture reused
  - Database structure reuse
  - Other large-scale design and code reuse
- *Cooperation Based*
  - Partners contributing processed data from their systems
  - Partners perhaps contributing "system-wide" tools
  - GHCC a part of the overall system
  - GV processing systems part of the overall system
  - Regional distribution centers part of the overall system
- Tools based approach to facilitate component isolation, extensibility and portability
  - Like TSDIS but more extensive
  - Multiple sources of tools anticipated
- Industry Standards based interactions
  - XML
  - Perhaps XML-SOAP based services interactions to facilitate
  - Facilitate contributor independences while encouraging broader interactions
- Distributed Science Discipline System (via NASA peer review, other initiatives) –
  - External entities at Universities or international locations
  - Focused on specific aspects of the overall precipitation research
  - Leading and acknowledged role in the precipitation community
  - Continuous and high speed, server level interactions with the Precipitation Processing System
  - Provide services and data products beyond the standard GPM products

Systematic Measurements from NPP



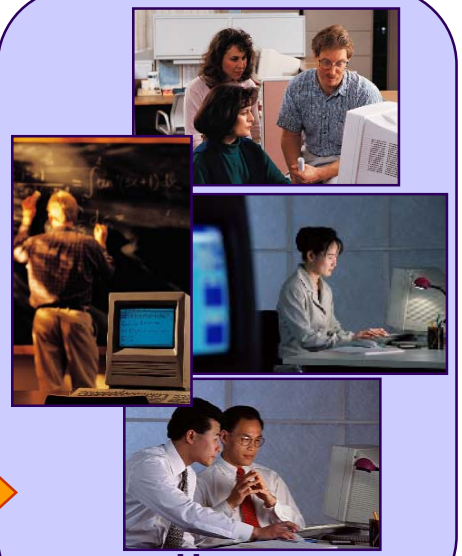
IPO
IDPS

Data Acquisition,
Data Capture,
Near-real-time
Processing


Weather Centrals

Info. Management,
Data Archive,
& Distribution

NOAA
CLASS


Users

Primary System

Instrument Development and Calibration/Validation

Heritage Measurement Series

Heritage Data Systems Refreshed/Expanded
Level 0/1 Not Predetermined

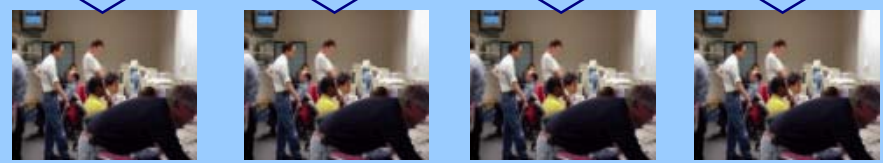
Ocean Color and SST | Terrestrial Suite | Atmospheric Suite | Cryospheric Products

- Size appropriate to research strategy
- Competed Science Team & IT support
- Civil Service "Project" leaders

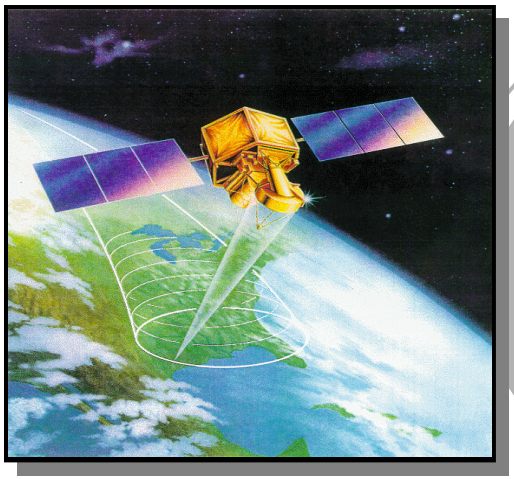
NPOESS
NPP


Science Team

Populates….


Systematic Measurement Teams

Long-term data assimilation feeds into climate models

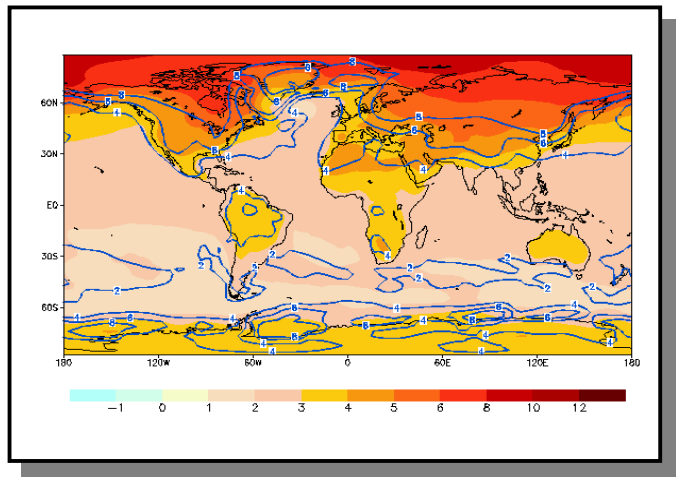


Long-term Observations

Biomass
Ocean
Carbon
Atmosphere
CO_2
Land
Aerosols
Clouds
Precipitation

Data assimilation, High-end climate modeling and computing

Statistics and analysis
Algorithms
Many Runs
Large Data Sets
Higher Resolution



- **Modeled climate forcings and feedbacks**
- **Projections of future climate states**
- **Global & Regional data product for assessments**

Managing the End-to-End Information Flow

Petabytes 10^{15}

Multi-platform, multi-parameter, high spatial and temporal resolution, remote & in-situ sensing

Calibration, Transformation To Characterized Geophysical Parameters

Terabytes 10^{12}

Interaction Between Modeling/Forecasting and Observation Systems

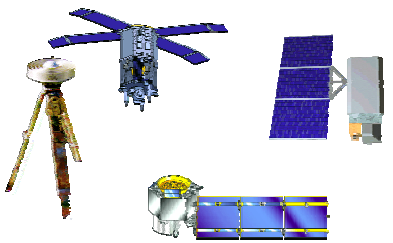
Gigabytes 10^{9}

Interactive Dissemination

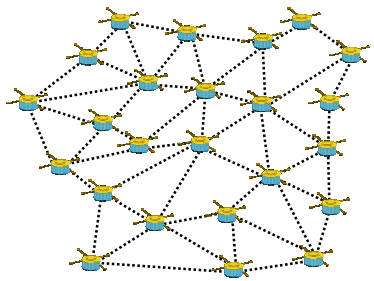
Predictions

Megabytes 10^{6}

Advanced Sensors

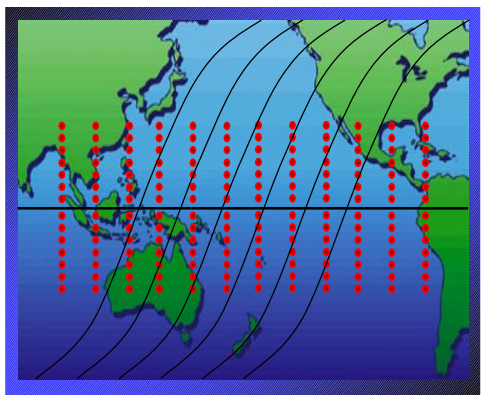


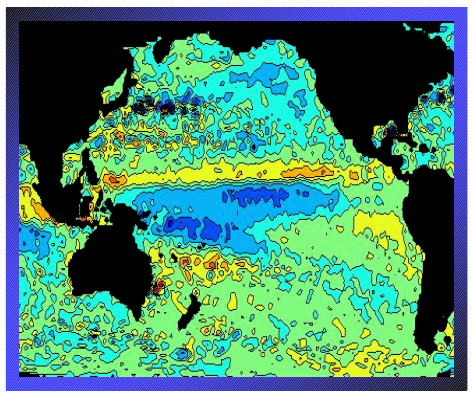
Data Processing & Analysis

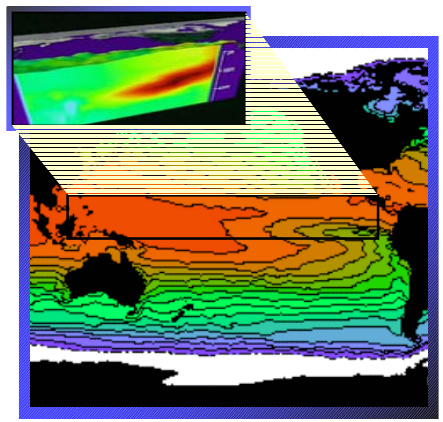


Information Synthesis

Access to Knowledge







Backup Charts

Research Solutions Solicited, 1

- **Projects Contributing to Systematic Missions**
  - ESE Research Strategy questions, esp. those concerning variability and forcing , need systematic measurements, i.e. accurate, uninterrupted series of key geophysical parameter records
  - Systematic data sets typically involve the synthesis of data from multiple instrument and/or platforms, and will exhibit consistent calibration and common validation throughout the entire measurement time series
  - Systematic data sets proposed should be highly useful to a significant segment of the Earth science research community in its efforts to help provide answers to the questions in the ESE Research Strategy, and also may be useful to a broader range of scientists and policy-makers in the context of assessment and support for environmental and policy decisions
  - Where there are different ideas in the research community about how to synthesize or choose among alternative algorithms for data sets, proposers should demonstrate that they will work with the community reach resolution and consensus to maximize acceptance and use.

Research Solutions Solicited, 2

- **Projects Contributing to Interdisciplinary or Process Studies**
  - Some science questions by their nature pose needs for concerted gathering of "bundles" of data, information and services:
    - Large regional scientific problems
    - Interdisciplinary scientific questions
    - Cycling questions
    - Large impact processes
  - Community-based aggregation of data, information, tools, and services dedicated to providing inputs to the problem at hand may be most effective
  - Information may need to be derived from disparate or multiple-source data with data-usage barriers (such as temporal and spatial differences) removed

CAN Examples for Research

Note: CAN scope is not limited to the areas contained in the examples listed.

- **Precipitation**
  - NASA moving to global precipitation after success with TRMM
  - Selected precipitation-related REASoN projects will work with SEEDS Prototype Integrated Precipitation Processing System

- **Data Assimilation Products Consistent for Climate Study**
  - Solicited climate data assimilation output data sets, which are consistent over time and space

- **Cryospheric Products**
  - Solicited cryospheric products to effectively address questions that pertain to ice on the land and the seas and their interaction with the surrounding oceans and atmosphere
  - The cryospheric products is a special opportunity